

Data driven performance improvement

Ben Thompson, January 2017

Have you ever changed your system because of a bad result? If so, how much was the system problem costing you?

Are you a better declarer than average? If you think so, then by how much?

If you're interested enough to be reading this, you're very likely amongst the top few percent of players in your bridge "field" – your state, your country, even the world. You know you're a pretty good player and you know your judgment is pretty good.

But just getting into the top 2% in **any** competitive endeavour won't win you any championships. Getting inside the top 1% won't get you on the national team. The vast majority of titles, and the spoils of those triumphs, go to a tiny fraction of competitors; barely one in a thousand **competitive** players, let alone recreational players.

Elite performers are ruthless about examining their own performance. They are obsessed with the detail. They are committed to continually making the tiny improvements that lift them out of the pack. If you're ruthless, obsessed and committed, read on.

Why do all that hard analysis?

It's very common for one big disaster or triumph to stick in the brain, and overwhelm all memory of similar situations with different (but usually smaller, more "normal") outcomes. It's a classic psychological trap and it really does work both ways.

Maybe your methods landed you in 5Cx in a 4-4 fit breaking 5-0 and going for 800 when everyone else was in 3NT but what about all the times you did the same thing and 5C made while 3NT went down? Maybe you opened 1H with an 8 count in 1st seat and talked the oppos out of a game but what about all the times partner believed you and you got too high? Without tracking the data and assessing the net result, you're flying blind if you try to adjust your system or your style.

Analysing the data forces you to be honest about your **whole** performance and enables you to make **informed** decisions about what to change and what to work on.

Beginnings of a Process

In all serious partnerships, I've always analysed our performance after a major tournament. Some partners have done the same thing and we've compared notes. Whether we'd done well or poorly, the objective was and is always to find opportunities for improvement.

Originally I did my number crunching with Excel, and I still use it a lot. I ordinarily classify hands into a variety of types (eg slam, game, defend doubled) and check if there's a weak performance area. Sometimes I subclassify hands (eg light games) to dig further into the detail. Sometimes something seems to have been a problem and I specifically subclassify hands to test if my/our *impression* matches *reality*.

I also often classify hands by opponent quality. If you aspire to perform well in serious company, you need to focus on how you perform in serious company.

Today, BBO records are excellent sources of play data, and there are tools that automatically do some basically hand classification for you, making it much easier to analyse performance.

Data

BBO play records are excellent data. Sometimes the operator misses the play, or even the auction, but by and large the quality is high and there's a **lot** of information in there. Having the whole auction means you can easily pull out hands where you played, or defended, or started 1S-2D, or whatever. The opening lead is almost always available (and correct!) and that's also good information.

Of course we all play many more hands than just those where we're on BBO. I usually dump online scorecard data into a spreadsheet, categorize the hands, and then I'm ready to crunch away. If I know going into a tournament that's there's something specific I want to look at, I annotate my scoresheet as I play to make sure I pick it up later when I'm categorizing the hands.

Measurement

How do you know if you've done well or poorly on a hand?

The way you measure your performance on a hand depends partly on what data you have to work with and partly on what you're interested in. All "measures" come with good features and poor ones; there's a tradeoff.

For IMP scoring, the most obvious measures are *actual* swing (teams) and *datum*. Even though there are problems with datums (to be brutal, a datum gives you the score a teammate you would never play with would bring back) but they're usually more useful than the actual swing because they factor out teammates randomly doing well or poorly on some board, which would unfairly skew your analysis of your own performance.

Datums aren't always available (eg the final) or are likely to be volatile (eg the semi-final). In that case, you could just give up and use the actual swing or you could try the *par* result. Par – the score achieved if everybody plays double-dummy perfect – is occasionally a wildly unrealistic measure but over a large number of hands it's "harsh but fair".

Unfortunately we all occasionally get ... mangled ... by the opponents doing something good or just plain lucky. That can feel like it would skew your results but those things tend to balance out, and actually sometimes you secretly contributed to the oppos doing well so it helps to keep that data exactly as is.

If you have the time, the inclination, and an unbiased mind, you could try constructing your own "theoretical datum". This can be quite helpful if you're looking only at your bidding performance. For example, if you're looking at how you go bidding light games and 4 of the 5 21 point games you bid went down to bad trump breaks, a simple number crunch might leave you thinking you're overbidding but a closer inspection of the hands would probably leave you comfortable with what you're doing. More on that later.

Tools

I use 3 tools to analyse performance.

Excel I've already talked about above. Formulas are great for aggregating and averaging data across your categories. If you're up to it, array formulas and pivot tables can be very helpful as well. If you're keen but not sure how to go about, you have an Excel power user somewhere in your life who

can't wait to help you out! I've provided an extract from a spreadsheet I did last year when I was looking at how our opening bids were performing.

Double Dummy Solver (DDS, from <http://www.bridgecaptain.com/downloadDD.html>) is an excellent tool for bulk analysis. For example, when Peter Hollands and Justin Howard made the 2017 Australian Open Team, I downloaded the lin file of every session they had played on BBO for the previous 3 years, loaded them all into DDS and crunched away. That informed some excellent conversations we had about what to think about and what to work on.

The last and most important tool is ... your Brain! Bald numerical analysis can lead to some very bad conclusions. The unlucky part of going down in 4 out 5 21 point games won't get picked up by Excel or DDS but when you look at those 5 hands, you'll see straight away that it was just an anti-percentage run and you'll be happy to keep on bidding 'em up.

Examples

Adjusting to new methods

Bill Jacobs and I played a system in which all of our 2 level openings were 10-13 and natural. That meant we dropped weak 2s. We kept track of our performance (in an Excel spreadsheet), including how we went when we passed an ordinary weak 2. It turned out that we were gaining plenty with the new openings, but we were dropping a fair bit on the weak 2s we were now passing. We were net ahead but we were dropping enough on the weak 2s to motivate a change.

The only realistic option was to option our weak 2s at the 3 level. We found quite quickly that our losses on those weak 2 hands dropped dramatically. Our detailed look suggested two things: (1) as you might expect, that we were losing a bit on the weak 2s relative to the field because we were opening a level too high; (2) that our performance on regular 3 level pre-empts dropped a bit (but not too much) because they had become looser. In the end we found it was fine to pass the poorer weak 2s, and our performance on weak 2 and weak 3 hands both improved a little.

Training, and planning for training

I mentioned above that I'm doing some work with Peter Hollands and Justin Howard, helping them to prepare for the 2017 Bermuda Bowl. Our opening conversation was around deciding where to focus, and crunching some numbers to see if we could identify some areas to work on was obvious.

Pete and Juzz are making a video series of their Journey to the Bowl and they talk about this opening analysis phase in the first video: https://www.youtube.com/watch?v=5h1PW_nfXnA

Here's some sample output from DDS:

Compared to double dummy (DD)	Actual Count	From Possible Deals	Percentage	Pct. over Opponent
We overbid/doubled	14	295	4.70%	0.60%
....missed slam	21	31	67.70%	18.30%
....missed game	30	114	26.30%	-4.90%
....missed best sacrifice	25	44	56.80%	-24.20%
They overbid/doubled	15	282	5.30%	
....missed slam	37	43	86.00%	
....missed game	24	112	21.40%	
....missed best sacrifice	14	43	32.60%	

The first (green) shaded line says that “we” (Pete & Juzz) missed bidding slam on 21 out of the 31 deals *where it was double dummy makeable* (ie 67.70% of the time). That seems bad at first glance but remember that double dummy making includes weird finesses, lucky drops and so on. When you’re thinking about double dummy performance, you need to compare your performance to some more realistic benchmark. DDS compares how you went double dummy compared to how everyone else with the same cards went double dummy. In this case, “we” went 18.30% better (last column) than “they” did.

So if we’re looking for something to work on, slam bidding probably isn’t it.

On the other hand, “we” seemed to miss the best sacrifice 24.20% more than “they”. Maybe that’s an area to work on.

Now is the time to dig deeper. There are lots of reasons why you might have missed the best sacrifice, and many of them are valid. So you have to start looking at hands in detail to understand what’s going on. Pete, Juzz and I looked at 17 game and slam hands where they missed a potential sacrifice. We concluded quickly that about half of them were too hard but that left half where we had a good discussion about how the auction might have gone differently, and a few methodological choices that might have helped.

DDS lets you filter your hand collection on all sorts of criteria (eg who opened what, who declares, point ranges) and you can look at both bidding and play performance.

[Next](#)

The key message is to use data analysis, not data points, to drive your decisions about system and practice. You can start anywhere, but I suggest you start simple to get a feel for what you can do, and work your way up.

And have a look at the attached article on Katie Ledecky, the best athlete on the planet. She is ruthless, obsessed and committed. Page 1 notes that she is routinely the best *in a skill that barely matters* simply because she approaches being the best as being the best at *everything*. Page 5 talks about how she eats up detailed performance analysis data to drive further improvement *even though she is already the best by a huge margin*.